

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/518,377

Source: Pg 7/10

Date Processed by STIC: 6/16/05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 06/16/2005

PATENT APPLICATION: US/10/518,377

TIME: 11:00:07

Input Set : A:\P26459.ST25.txt

Output Set: N:\CRF4\06162005\J518377.raw

```

3 <110> APPLICANT: KANG, Hyun-Ah
4     RHEE, Sang-Ki
5     SOHN, Min-Jeong
6     KIM, Jeong-Yoon
8 <120> TITLE OF INVENTION: HANSENULA POLYMORPHA YAPSIN DEFICIENT MUTANT STRAIN AND
PROCESS
9     FOR THE PREPARATION OF THE RECOMBINANT PROTEINS USING THE SAME
11 <130> FILE REFERENCE: P26459
13 <140> CURRENT APPLICATION NUMBER: US 10/518,377
14 <141> CURRENT FILING DATE: 2004-12-28
16 <150> PRIOR APPLICATION NUMBER: PCT/KR2003/001279
17 <151> PRIOR FILING DATE: 2003-06-28
19 <160> NUMBER OF SEQ ID NOS: 16
21 <170> SOFTWARE: PatentIn version 3.3
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 3151
25 <212> TYPE: DNA
26 <213> ORGANISM: Hansenula polymorpha
28 <400> SEQUENCE: 1
29 agttgagtcg caatagtggt gcgaacttca aatgccctta ctgtccgcga acaaccacca      60
31 ttgccaggc tgtgcaggcc agatttggtta atttgtgaaa agtggaaaaa atttattccg      120
33 ctatgcctaa ccgaagagcc cgcaagaaga ggccgacaga agacttttcc agctcttcgg      180
35 catctgaaaa cgatagtgac tccgagagcg tgaccagtgt acaggaagag cagccggatg      240
37 cgcccgaaac atacacaata gatggcctgg acacgcaaga ggtgtctgac agcacacagg      300
39 tgagactcca acagctgaac gcagacaggt tggccagcat agagcaaagc ctttcaggca      360
41 acctcaaaact ggacataaac gcagtagcgc agatagatga tgtgcgtgag cagctgcaga      420
43 acgagtattt gaagaaattg cttgtcacat attctgagga cctggatgag ctgcgtcaga      480
45 aaaccgattt caaggaaaac tcactcaaaa ccctcgcccg tcttctcaaa gagagcggaa      540
47 acatatttga tgatggaact ctcaagtcgc tagttgagtg atgtatatga taatgtctaa      600
49 ttttaatttt catcagtggt caagatctgg gcttagccgt tctaaatggg atattcaggc      660
51 tgtgcaagcc acatttaaaa ttaccccatc ggtttttaa tttctattgt agaaattagg      720
53 atctacatag aggtagagtg agcaacagaa cattgtttgc tatccggggc ctccgactgg      780
55 aacgtcttac cttcagctac tatttattca gaaaaaagag tgcattttca tctatcaagg      840
57 tctcaaaagt tcgaatcaaa tcactagtat tttttccgag actaaaaaaa agttgacaca      900
59 atgaaagttg ctacactgtt tttcttggtc tcgagtgtct gtgtgctggg agaccacag      960
61 ttcgtgaaac tggaggcctc tgttcttcgg ggatccactt acaaggattc ccagaagggg      1020
63 gccaaaccgt tcatgttga aaagagggtc gatgacggct cggtcacgat ggaattgcag      1080
65 aacgcccagt ctttctacca agtcgagatc gagataggat ctgataagca gaaggtgggg      1140
67 gttttgattg ataccggttc ctcggacttg tgggtgatga actcgaataa ctcttactgt      1200
69 tcgtcttcca gcactaaaaa attgaaacgg gacggaccgg ccgatgctgt acaaaaagga      1260
71 cgcgatcttt ccgacctgta caatttcaac tctccaaacg aagacaacaa tgcaaaaagga      1320
73 ttcttgggtg gctggggaga cttgacaca gttagactg caaccaggga tgagacacag      1380
75 acggctctcg ctgcgcaggc caccgtggac tgctcgctat acggaacggt caatccttca      1440
77 acgtccaatt cgttccacaa caacggcacc acatttgaga ttctgtacgc ggaccgcact      1500

```

RAW SEQUENCE LISTING

DATE: 06/16/2005

PATENT APPLICATION: US/10/518,377

TIME: 11:00:07

Input Set : A:\P26459.ST25.txt

Output Set: N:\CRF4\06162005\J518377.raw

```

79 tttgcccgtg gaacctgggg ctacgatgat gtcactttca atggtgtcac ggtaaacgat 1560
81 ctctcgttgg ccgtggcaga tgaacacagat tcttcgactg gtgtttttgg tatcggattg 1620
83 aggggaattgg aaaccacata ctacaggaggc ggaccacagc attacatcta cgacaactta 1680
85 cttttcaaaa tggctcgacca gggactcatc aatagagccg cctattccgt ctacctgaac 1740
87 tcaactgagt ccagcactgc ctcgatcctc ttcggtgcgg ttgaccaaag caaatatacc 1800
89 ggaagtcttg gcttgcttcc tatcatcaac acggctgctt cctacggtta ccaaaagcct 1860
91 ctaaggtccc aaatcaccct gtctgccatt acggtcagcg actccagagg acagcaagca 1920
93 agcattgggtt caggagctgc tgctgcactt cttgataccg gaacgacttt gacgtatgct 1980
95 ccaagcgaga ttgtcgagaa acttgctgaa accctaggct tcgactacag cagctctgtc 2040
97 ggggcctacg tggcaagatg cagggacggt gatagctacg ctgtcaactt cgacttccag 2100
99 ggtaaagtga ttgaagctcc tttgagttcc ttcctgattg ctctgcaaac caactccgga 2160
101 gaagtttcct cctactgcgc attgggtatt ttctcctctg gagacgaatc cttcacgctc 2220
103 ggcgatactt tctgcgaaa cgcctacttt gtggctgacc tcgagggata tcaaactcgt 2280
105 atagctaacy tgaacctgaa tcttgagacc gagcaaatg aggtcatctc aggcactcc 2340
107 attccttctg cttcgtcggg ttccgattac tccaatacct ggggcgcctc tgccaccgct 2400
109 ttggacactg acaggcctac tactctggga tctgtgactg ctgtgggcca tgaaagagtg 2460
111 acctcgacca agaaggtttc gagtgtgaag acaagcactt cgtccggggtc cgggtccact 2520
113 tcggagtcgt ctacgtccag ttcgcattcc agcaatggcc caaggacagt aggccttagt 2580
115 ttgtgtgccg ttttgtgcgc attcttgatt tctatactag ttgtttgcta gatctgaagt 2640
117 tctaaggggc tttagtcttc atttatgatt tttttttatt tggaccgcct cgaattgttt 2700
119 ttccgacggg tctactttaa agctgcaaga tctcgtttag cgtcgtttat ttctcgttcg 2760
121 tttagtgaac aaaaaacaga aaaaaaaact ataaaaagcg gtatataacc tttatatttt 2820
123 gataaacatg agcagcgaaa ttaagctagc accaaaggat tacgagaagg acaaggagtt 2880
125 cgccaaggct ctgcatggca aggacgccgc gagcgctaca ggaatgagtg cttgggtgaa 2940
127 gaaggacaag gaagctcaaa aagtcgcgat ggaaggatat ttcaagcact gggacgggaa 3000
129 aaccgacgag gagactgaaa agtcgagact cgaggactac tcgacgctca ccaagcacta 3060
131 ctacaacctg gtgacggatt tctacgagta tggatgggga tcctcgttcc acttttccag 3120
133 atactacaag ggagagccat ttagacaagc t 3151

```

136 <210> SEQ ID NO: 2

137 <211> LENGTH: 576

138 <212> TYPE: PRT

139 <213> ORGANISM: Hansenula polymorpha

141 <400> SEQUENCE: 2

```

143 Met Lys Val Ala Thr Leu Phe Phe Leu Ala Ser Ser Val Cys Val Leu
144 1 5 10 15
147 Gly Asp Pro Gln Phe Val Lys Leu Glu Ala Ser Val Leu Arg Gly Ser
148 20 25 30
151 Thr Tyr Lys Asp Ser Gln Lys Gly Ala Lys Pro Phe Met Leu Glu Lys
152 35 40 45
155 Arg Ala Asp Asp Gly Ser Val Thr Met Glu Leu Gln Asn Ala Gln Ser
156 50 55 60
159 Phe Tyr Gln Val Glu Ile Glu Ile Gly Ser Asp Lys Gln Lys Val Gly
160 65 70 75 80
163 Val Leu Ile Asp Thr Gly Ser Ser Asp Leu Trp Val Met Asn Ser Asn
164 85 90 95
167 Asn Ser Tyr Cys Ser Ser Ser Ser Thr Lys Lys Leu Lys Arg Asp Gly
168 100 105 110
171 Pro Ala Asp Ala Leu Gln Lys Gly Arg Asp Leu Ser Asp Leu Tyr Asn
172 115 120 125

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/518,377

DATE: 06/16/2005

TIME: 11:00:07

Input Set : A:\P26459.ST25.txt

Output Set: N:\CRF4\06162005\J518377.raw

```

175 Phe Asn Ser Pro Asn Glu Asp Asn Asn Ala Lys Gly Phe Leu Gly Gly
176      130      135      140
179 Trp Gly Asp Leu Thr Thr Val Glu Thr Ala Thr Gln Asp Glu Thr Gln
180 145      150      155      160
183 Thr Ala Leu Ala Ala Gln Ala Thr Val Asp Cys Ser Leu Tyr Gly Thr
184      165      170      175
187 Phe Asn Pro Ser Thr Ser Asn Ser Phe His Asn Asn Gly Thr Thr Phe
188      180      185      190
191 Glu Ile Ser Tyr Ala Asp Arg Thr Phe Ala Arg Gly Thr Trp Gly Tyr
192      195      200      205
195 Asp Asp Val Thr Phe Asn Gly Val Thr Val Asn Asp Leu Ser Leu Ala
196      210      215      220
199 Val Ala Asp Glu Thr Asp Ser Ser Thr Gly Val Phe Gly Ile Gly Leu
200 225      230      235      240
203 Arg Glu Leu Glu Thr Thr Tyr Ser Gly Gly Pro Gln His Tyr Ile
204      245      250      255
207 Tyr Asp Asn Leu Pro Phe Lys Met Val Asp Gln Gly Leu Ile Asn Arg
208      260      265      270
211 Ala Ala Tyr Ser Val Tyr Leu Asn Ser Thr Glu Ser Ser Thr Ala Ser
212      275      280      285
215 Ile Leu Phe Gly Ala Val Asp Gln Ser Lys Tyr Thr Gly Ser Leu Gly
216      290      295      300
219 Leu Leu Pro Ile Ile Asn Thr Ala Ala Ser Tyr Gly Tyr Gln Lys Pro
220 305      310      315      320
223 Leu Arg Leu Gln Ile Thr Leu Ser Ala Ile Thr Val Ser Asp Ser Arg
224      325      330      335
227 Gly Gln Gln Ala Ser Ile Gly Ser Gly Ala Ala Ala Ala Leu Leu Asp
228      340      345      350
231 Thr Gly Thr Thr Leu Thr Tyr Ala Pro Ser Glu Ile Val Glu Lys Leu
232      355      360      365
235 Ala Glu Thr Leu Gly Phe Asp Tyr Ser Ser Ser Val Gly Ala Tyr Val
236      370      375      380
239 Ala Arg Cys Arg Asp Val Asp Ser Tyr Ala Val Asn Phe Asp Phe Gln
240 385      390      395      400
243 Gly Lys Val Ile Glu Ala Pro Leu Ser Ser Phe Leu Ile Ala Leu Gln
244      405      410      415
247 Thr Asn Ser Gly Glu Val Ser Ser Tyr Cys Ala Leu Gly Ile Phe Ser
248      420      425      430
251 Ser Gly Asp Glu Ser Phe Thr Leu Gly Asp Thr Phe Leu Arg Asn Ala
252      435      440      445
255 Tyr Phe Val Ala Asp Leu Glu Gly Tyr Gln Ile Ala Ile Ala Asn Val
256      450      455      460
259 Asn Leu Asn Pro Gly Ala Glu Gln Ile Glu Val Ile Ser Gly Asn Ser
260 465      470      475      480
263 Ile Pro Ser Ala Ser Ser Val Ser Asp Tyr Ser Asn Thr Trp Gly Ala
264      485      490      495
267 Ser Ala Thr Ala Leu Asp Thr Asp Arg Pro Thr Thr Leu Gly Ser Val
268      500      505      510
271 Thr Ala Val Gly Asp Glu Arg Val Thr Ser Thr Lys Lys Val Ser Ser

```

RAW SEQUENCE LISTING

DATE: 06/16/2005

PATENT APPLICATION: US/10/518,377

TIME: 11:00:07

Input Set : A:\P26459.ST25.txt

Output Set: N:\CRF4\06162005\J518377.raw

```

272          515          520          525
275 Val Lys Thr Ser Thr Ser Ser Gly Ser Gly Ser Thr Ser Glu Ser Ser
276          530          535          540
279 Thr Ser Ser Ser His Ser Ser Asn Gly Pro Arg Thr Val Gly Phe Ser
280 545          550          555          560
283 Leu Cys Ala Val Leu Cys Ala Phe Leu Ile Ser Ile Leu Val Val Cys
284          565          570          575
287 <210> SEQ ID NO: 3
288 <211> LENGTH: 25
289 <212> TYPE: DNA
290 <213> ORGANISM: Artificial Sequence
292 <220> FEATURE:
293 <223> OTHER INFORMATION: primer
295 <400> SEQUENCE: 3
296 gaagtgcagc agcagctcct gaacc
299 <210> SEQ ID NO: 4
300 <211> LENGTH: 26
301 <212> TYPE: DNA
302 <213> ORGANISM: Artificial Sequence
304 <220> FEATURE:
305 <223> OTHER INFORMATION: primer
307 <400> SEQUENCE: 4
308 ggctgatgac ggctcgggtca cgatgg
311 <210> SEQ ID NO: 5
312 <211> LENGTH: 20
313 <212> TYPE: DNA
314 <213> ORGANISM: Artificial Sequence
316 <220> FEATURE:
317 <223> OTHER INFORMATION: primer
319 <400> SEQUENCE: 5
320 ggacacgcaa gaggtgtctg
323 <210> SEQ ID NO: 6
324 <211> LENGTH: 40
325 <212> TYPE: DNA
326 <213> ORGANISM: Artificial Sequence
328 <220> FEATURE:
329 <223> OTHER INFORMATION: primer
331 <400> SEQUENCE: 6
332 agctcgctac ccgggggatcc gcaactttca ttgtgtcaac
335 <210> SEQ ID NO: 7
336 <211> LENGTH: 40
337 <212> TYPE: DNA
338 <213> ORGANISM: Artificial Sequence
340 <220> FEATURE:
341 <223> OTHER INFORMATION: primer
343 <400> SEQUENCE: 7
344 gcacatcccc ctttcgccag cctcttcggt gcggttgacc
347 <210> SEQ ID NO: 8
348 <211> LENGTH: 20

```

RAW SEQUENCE LISTING

DATE: 06/16/2005

PATENT APPLICATION: US/10/518,377

TIME: 11:00:07

Input Set : A:\P26459.ST25.txt

Output Set: N:\CRF4\06162005\J518377.raw

```

349 <212> TYPE: DNA
350 <213> ORGANISM: Artificial Sequence
352 <220> FEATURE:
353 <223> OTHER INFORMATION: primer
355 <400> SEQUENCE: 8
356 gctcggctcc aggattcagg                20
359 <210> SEQ ID NO: 9
360 <211> LENGTH: 20
361 <212> TYPE: DNA
362 <213> ORGANISM: Artificial Sequence
364 <220> FEATURE:
365 <223> OTHER INFORMATION: primer
367 <400> SEQUENCE: 9
368 ggatccccgg gtaccgagct                20
371 <210> SEQ ID NO: 10
372 <211> LENGTH: 20
373 <212> TYPE: DNA
374 <213> ORGANISM: Artificial Sequence
376 <220> FEATURE:
377 <223> OTHER INFORMATION: primer
379 <400> SEQUENCE: 10
380 caccggtagc taatgatccc                20
383 <210> SEQ ID NO: 11
384 <211> LENGTH: 20
385 <212> TYPE: DNA
386 <213> ORGANISM: Artificial Sequence
388 <220> FEATURE:
389 <223> OTHER INFORMATION: primer
391 <400> SEQUENCE: 11
392 cgaacatcca agtgggcccga                20
395 <210> SEQ ID NO: 12
396 <211> LENGTH: 20
397 <212> TYPE: DNA
398 <213> ORGANISM: Artificial Sequence
400 <220> FEATURE:
401 <223> OTHER INFORMATION: primer
403 <400> SEQUENCE: 12
404 ctggcgaaag ggggatgtgc                20
407 <210> SEQ ID NO: 13
408 <211> LENGTH: 24
409 <212> TYPE: DNA
410 <213> ORGANISM: Artificial Sequence
412 <220> FEATURE:
413 <223> OTHER INFORMATION: primer
415 <400> SEQUENCE: 13
416 gaattcatga agtgggtaac cttt          24
419 <210> SEQ ID NO: 14
420 <211> LENGTH: 20
421 <212> TYPE: DNA

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/518,377

DATE: 06/16/2005

TIME: 11:00:08

Input Set : A:\P26459.ST25.txt

Output Set: N:\CRF4\06162005\J518377.raw